Type I/II errors and Exam Practice

Shannon Pileggi

STAT 217

OUTLINE

When we draw a conclusion about a hypothesis test (ie, reject or fail to reject H_0), do we always make the correct decision?

1. Yes
2. No

Possible outcomes of a hypothesis test

Decision based on observed data	**Unknown Truth**	
	H_0 true	H_0 false
Fail to reject H_0		
Reject H_0		

Possible outcomes of a hypothesis test

Decision based on observed data	**Unknown Truth**	
	H_0 true	H_0 false
Fail to reject H_0	Correct Decision	
Reject H_0		

Possible outcomes of a hypothesis test

Decision based on observed data	**Unknown Truth**	
	H_0 true	H_0 false
Fail to reject H_0	Correct Decision	
Reject H_0		Correct Decision

Possible outcomes of a hypothesis test

Decision based on observed data	**Unknown Truth**	
	H_0 true	H_0 false
Fail to reject H_0	Correct Decision	
Reject H_0	Type I Error	Correct Decision

- A **Type I error** occurs when H_0 is true in reality but is rejected based on evidence from the test.

Possible outcomes of a hypothesis test

Decision based on observed data	Unknown Truth: H_0 true	Unknown Truth: H_0 false
Fail to reject H_0	Correct Decision	Type II Error
Reject H_0	Type I Error	Correct Decision

- A **Type I error** occurs when H_0 is true in reality but is rejected based on evidence from the test.
- A **Type II error** occurs when H_0 is false in reality (or H_a true) but you fail to reject H_0 based on evidence from the test.

Type I/Type II error example

Jury trial

Decision by the jury	**Unknown Truth**	
	Defendant is innocent (H_0)	Defendant is guilty (H_a)
Innocent		
Guilty		

Type I/Type II error example

Jury trial

	Unknown Truth	
Decision by the jury	Defendant is innocent (H_0)	Defendant is guilty (H_a)
Innocent	Correct Decision	
Guilty		Correct Decision

Type I/Type II error example

Jury trial

Decision by the jury	Unknown Truth: Defendant is innocent (H_0)	Unknown Truth: Defendant is guilty (H_a)
Innocent	Correct Decision	
Guilty	Type I Error	Correct Decision

- A **Type I error** occurs when the jury finds a truly innocent man to be guilty.

Type I/Type II error example

Jury trial

Decision by the jury	**Unknown Truth**	
	Defendant is innocent (H_0)	Defendant is guilty (H_a)
Innocent	Correct Decision	Type II Error
Guilty	Type I Error	Correct Decision

- A **Type I error** occurs when the jury finds a truly innocent man to be guilty.
- A **Type II error** occurs when the jury finds a truly guilty man to be innocent.

Type I/Type II error example

Jury trial

Decision by the jury	Unknown Truth: Defendant is innocent (H_0)	Unknown Truth: Defendant is guilty (H_a)
Innocent	Correct Decision	Type II Error
Guilty	Type I Error	Correct Decision

- A **Type I error** occurs when the jury finds a truly innocent man to be guilty.
- A **Type II error** occurs when the jury finds a truly guilty man to be innocent.

Which error do you think is worse?

1. Type I
2. Type II

Type I/Type II error example

	Unknown Truth	
Decision based on screening test	Woman does not have breast cancer (H_0)	Woman has breast cancer (H_a)
No suspicion of breast cancer	Correct Decision	Type II Error
Suspicion of breast cancer	Type I Error	Correct Decision

Type I/Type II error example

Decision based on screening test	Unknown Truth: Woman does not have breast cancer (H_0)	Woman has breast cancer (H_a)
No suspicion of breast cancer	Correct Decision	Type II Error
Suspicion of breast cancer	Type I Error	Correct Decision

- A **Type I error** occurs when the woman does not have breast cancer, but the mammogram indicates that she may. This is a false positive.

Type I/Type II error example

Decision based on screening test	Unknown Truth: Woman does not have breast cancer (H_0)	Unknown Truth: Woman has breast cancer (H_a)
No suspicion of breast cancer	Correct Decision	Type II Error
Suspicion of breast cancer	Type I Error	Correct Decision

- A **Type I error** occurs when the woman does not have breast cancer, but the mammogram indicates that she may. This is a false positive.
- A **Type II error** occurs when the woman does have breast cancer, but the mammogram indicates that does not. This is a false negative.

Type I/Type II error example

Decision based on screening test	Unknown Truth: Woman does not have breast cancer (H_0)	Unknown Truth: Woman has breast cancer (H_a)
No suspicion of breast cancer	Correct Decision	Type II Error
Suspicion of breast cancer	Type I Error	Correct Decision

- A **Type I error** occurs when the woman does not have breast cancer, but the mammogram indicates that she may. This is a false positive.
- A **Type II error** occurs when the woman does have breast cancer, but the mammogram indicates that does not. This is a false negative.

Which error do you think is worse?

1. Type I
2. Type II

If I conduct a hypothesis test and I reject the null hypothesis based on evidence from my data, this could have been the result of... *(mark all that apply)*

1. a Type I error
2. a Type II error
3. a correct decision

Type I error

- A **Type I error** occurs when H_0 is true but is rejected.
- $\alpha = \Pr(\text{Type I error}) = \Pr(\text{reject } H_0 \text{ when } H_0 \text{ true})$
- The choice of significance level controls the probability of a Type I error.
- The more serious the consequences of a Type I error, the smaller α should be.

Type II error

- A **Type II error** occurs when H_0 is false (or H_a true) but H_0 is *not* rejected.

  Pr(Type II error)=Pr(fail to reject H_0 when H_0 false)
- The Pr(Type II error) is a function of many things, including: α, n, s, and μ_0 - you won't need to calculate this.

Suppose we interested in determining if the population mean length of the longest serious relationship among Cal Poly students differs from 9 months. That is, we test H_0: $\mu = 9$ vs H_a: $\mu \neq 9$.

What would be an example of a Type I error? Finding that we (do/do not) have evidence that the population mean length of longest serious relationship differs from 9 when in reality the population mean length of longest serious relationship (equals/does not equal) 9 months.

1. do; equals
2. do not; does not equal
3. do; does not equal
4. do not; equals

We can't have it all

- We cannot simultaneously minimize Pr(Type I error) and Pr(Type II error).
- Pr(Type I error) and Pr(Type II error) are inversely related - as one goes down the other must go up.
- If we minimize the chance of a Type I error by making α smaller, we increase our chances of committing a Type II error.

Importance of conditions

We always have some conditions we need for *valid* inference. When these conditions are violated, but we use the statistical method anyway, this results in *invalid* inference. Invalid inference can mean:

- your confidence interval isn't capturing the parameter as often as it should (ie, lower than 95%)
- you commit Type I errors more than you should (ie, more than 5%)
- basically, can't trust your CI or your p-value

Errors in Hypothesis Testing

Practice Problems

For which of the following p-values would I reject the null hypothesis at the $\alpha = 0.10$ level of significance? Mark all that apply.

1. 0.01
2. 0.08
3. 0.15
4. 0.36
5. 0.89

Clicker

We are interested in describing the math SAT scores of Cal Poly students. Assume that math SAT scores of Cal Poly students are right-skewed with $\mu = 600$ and $\sigma = 75$. Consider the sampling distribution of the sample mean when we take a sample of size $n = 100$.

Can we assume that the sampling distribution of the sample mean is approximately normally distributed?

1. Yes, because n is bigger than 30.
2. Yes, because n is bigger than 10.
3. Yes, because $np > 10$ and $n(1 - p) > 10$.
4. No, because np or $n(1 - p)$ is not greater than 10.
5. No, because the population is right-skewed.

Below are the summary statistics of the data and output from the analysis testing if the population average birth weight of the monkeys is 0.4kg.

```
  min   Q1 median  Q3  max mean   sd  n missing
 0.27 0.37   0.39 0.5 0.68 0.44 0.12 10       0

t = 1.0853, df = 9, p-value = 0.306
alternative hypothesis: true mean is not equal to 0.4
95 percent confidence interval:
 XXXXXXX XXXXXXX
```

Which is the correct calculation to estimate the population average birth weight of rhesus monkeys with a 95% CI?

1. $0.44 \pm 1.0853 \times 0.12/\sqrt{10}$
2. $0.44 \pm 1.0853 \times 0.12$
3. $0.44 \pm 2.26 \times 0.12/\sqrt{10}$
4. $0.39 \pm 1.96 \times 0.12$
5. $0.39 \pm 2.26 \times 0.12/\sqrt{9}$

Suppose I test if the average age of retirement for Americans differs from 65 (H_0: $\mu = 65$ vs H_a: $\mu \neq 65$), and I get a test statistic of 1.5. Which figure corresponds to the p-value?

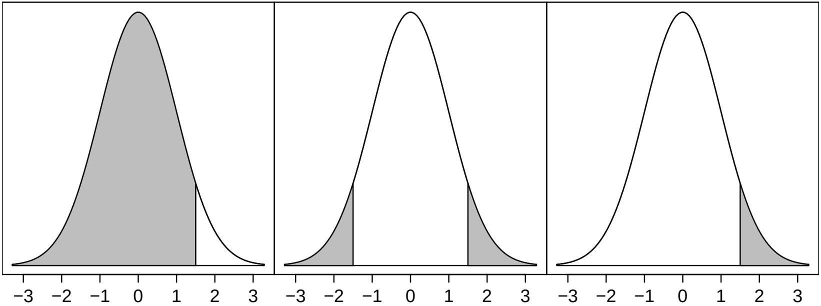


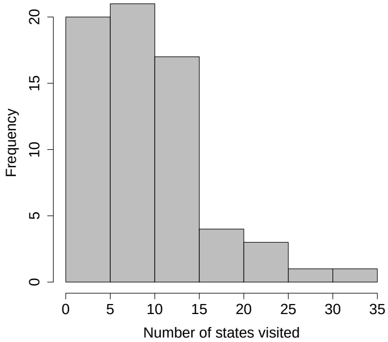
Figure 1 Figure 2 Figure 3

The level of significance α at which the test is performed (and hence, the corresponding confidence level of a CI) affects...

1. both the value of the test statistic and the width of a confidence interval.
2. only the value of the test statistic and not the width of a confidence interval.
3. only the width of a confidence interval and not the value of the test statistic.
4. neither the value of the test statistic nor the width of a confidence interval.

Number of states visited

```
> favstats(survey$States)
 min Q1 median Q3 max     mean       sd  n missing
   1  5      8 13  35 9.820896 6.797783 67       0
```



Are conditions satisfied for the following one-sample t-test? H_0: $\mu = 8$ vs H_a: $\mu \neq 8$, where $\mu =$ population mean number of states visited

1. yes, $n \geq 30$
2. no, distribution is right skewed
3. yes, $n\hat{p} \geq 10$ and $n(1-\hat{p}) \geq 10$
4. no, $n\hat{p} < 10$ OR $n(1-\hat{p}) < 10$

```
        One Sample t-test

data:  survey$States
t = 2.1926, df = 66, p-value = 0.03187
alternative hypothesis: true mean is not equal to 8
95 percent confidence interval:
  8.162786 11.479005
sample estimates:
mean of x
 9.820896
```

Which of the following statements are true at $\alpha = 0.05$?

1. 67 individuals were included in this sample
2. the null hypothesis is $\mu \neq 8$
3. the z distribution was used to calculate the p-value
4. there is evidence that the null hypothesis is true
5. all are true
6. none are true

Suppose Cal Poly administrators want to determine the proportion of students that graduate within 4 years, and they calculate a 95% confidence interval to be (0.4, 0.8). They aren't happy with the interval because it is so wide, and hence not very informative. What would you recommend that they do?

1. increase their confidence level
2. take a larger sample
3. perform a hypothesis test
4. reduce the standard deviation of whether or not students graduate
5. more than one of these items

The percentage of Cal Poly students that participate in Greek life is 20%. Suppose I were to repeatedly sample groups of $n = 100$ Cal Poly students. Describe the sampling distribution of the sample proportion in terms of the

- shape
- mean
- standard deviation

of the distribution.

Clicker

We tested if the average amount of hours worked per week differs from 38 at the $\alpha = 0.01$ level of significance. (H_0: $\mu = 38$ vs H_a: $\mu \neq 38$) In the data, $\bar{x} = 37.1$, and we get a p-value equal to 0.0299.

What can be said about a 99% confidence interval for the population mean?

1. A 99% confidence interval would contain 38.
2. A 99% confidence interval would not contain 38.
3. A 99% confidence interval would not contain 37.1.
4. A 99% confidence interval would contain 0.0299.
5. A 99% confidence interval would contain 0.
6. Not enough information to determine.

Clicker

Supposed I used a one-sample t-test to determine if average income of US citizens is less than \$45,000.
H_0: $\mu = 45000$ vs H_a: $\mu < 45000$

What would a Type II error be in the context of this research question? Finding that we (do/do not) have evidence that the population average income is less than \$45,000, when in reality the population average income (equals \$45,000/is less than \$45,000).

1. do not; is less than \$45,000
2. do not; equals \$45,000
3. do; equals \$45,000
4. do; is less than \$45,000